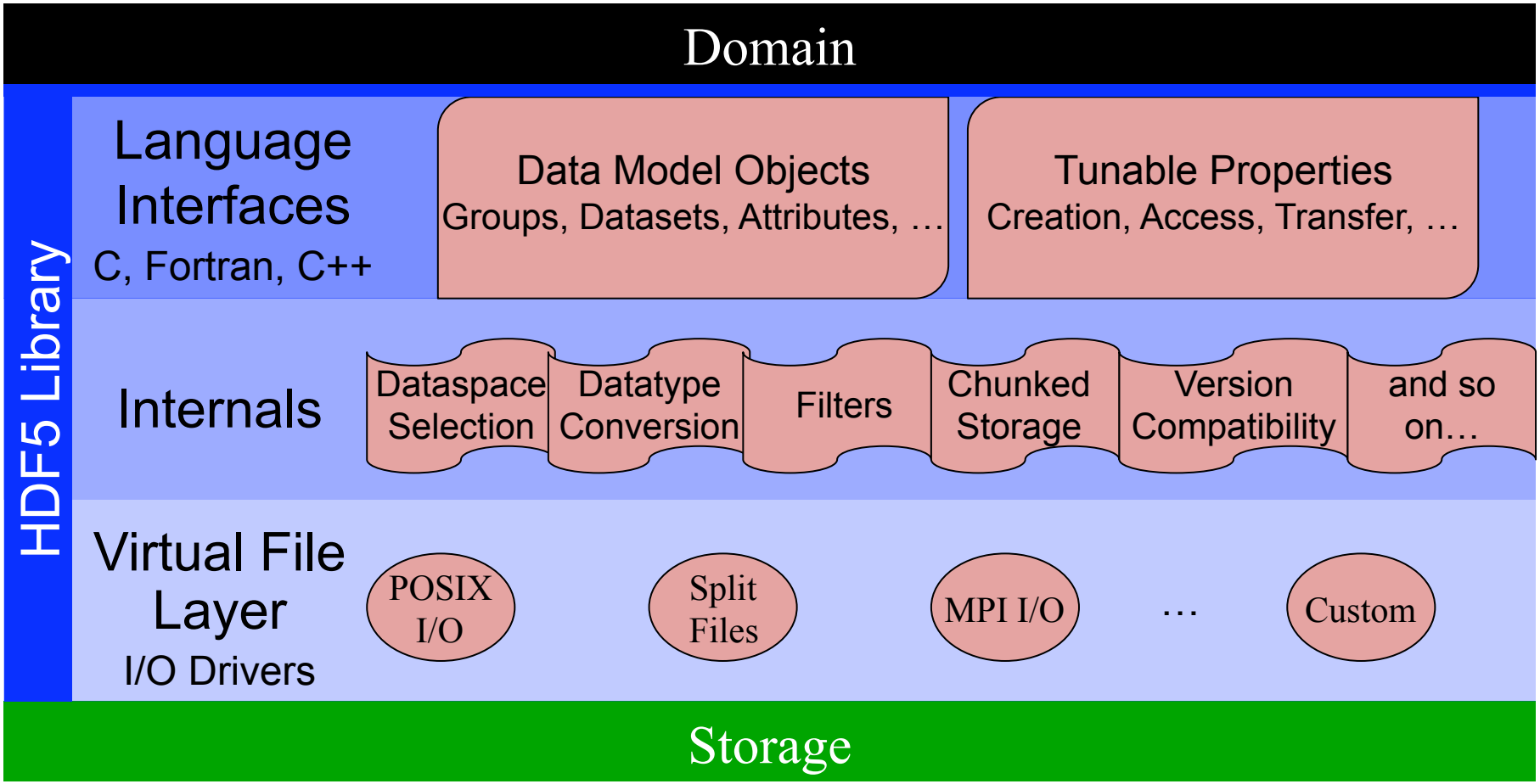
Tuning Parallel HDF5 for High Performance Computing Applications

Quincey Koziol
The HDF Group
koziol@hdfgroup.org

HDF5 Technologies

- HDF5 Abstract Data Model
  - Groups, Datasets, Attributes, …
- HDF5 Software
  - Tools
  - High-level Libraries
  - Fortran, C++, Java Wrappers
  - HDF5 C Library
- HDF5 Binary File Format
  - Bit-level organization of stored data

HDF5 Library Layers

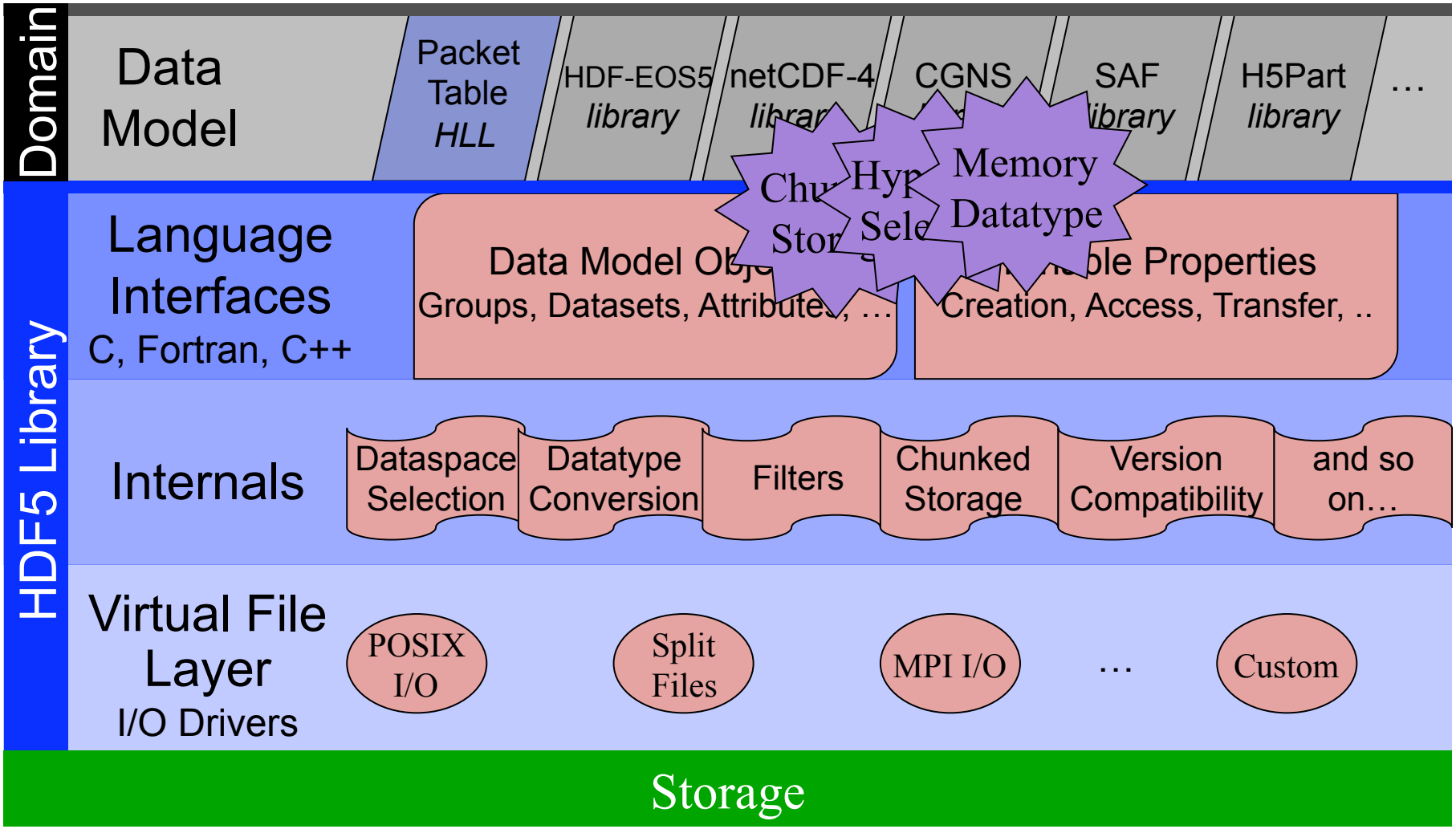


Domain Data Models in HDF5

- Every domain has a set of "concepts"
  - CFD: grid coordinates, boundary conditions, solutions, …
  - Test Data: calibrations, units, measurements, timestamps, …
  - Bioinformatics: genes, sequences, exons, pcrs, amplicons, …

- Map domain concepts to HDF5 abstract data model
  - Represent "concepts" as H5 Groups, Datasets, Attributes, …
  - Conventions supply semantics
  - Choose among many possible mappings
    - Lots of groups or just a few? Compound datasets? Variable Length data? …
    - *Choices will affect performance*

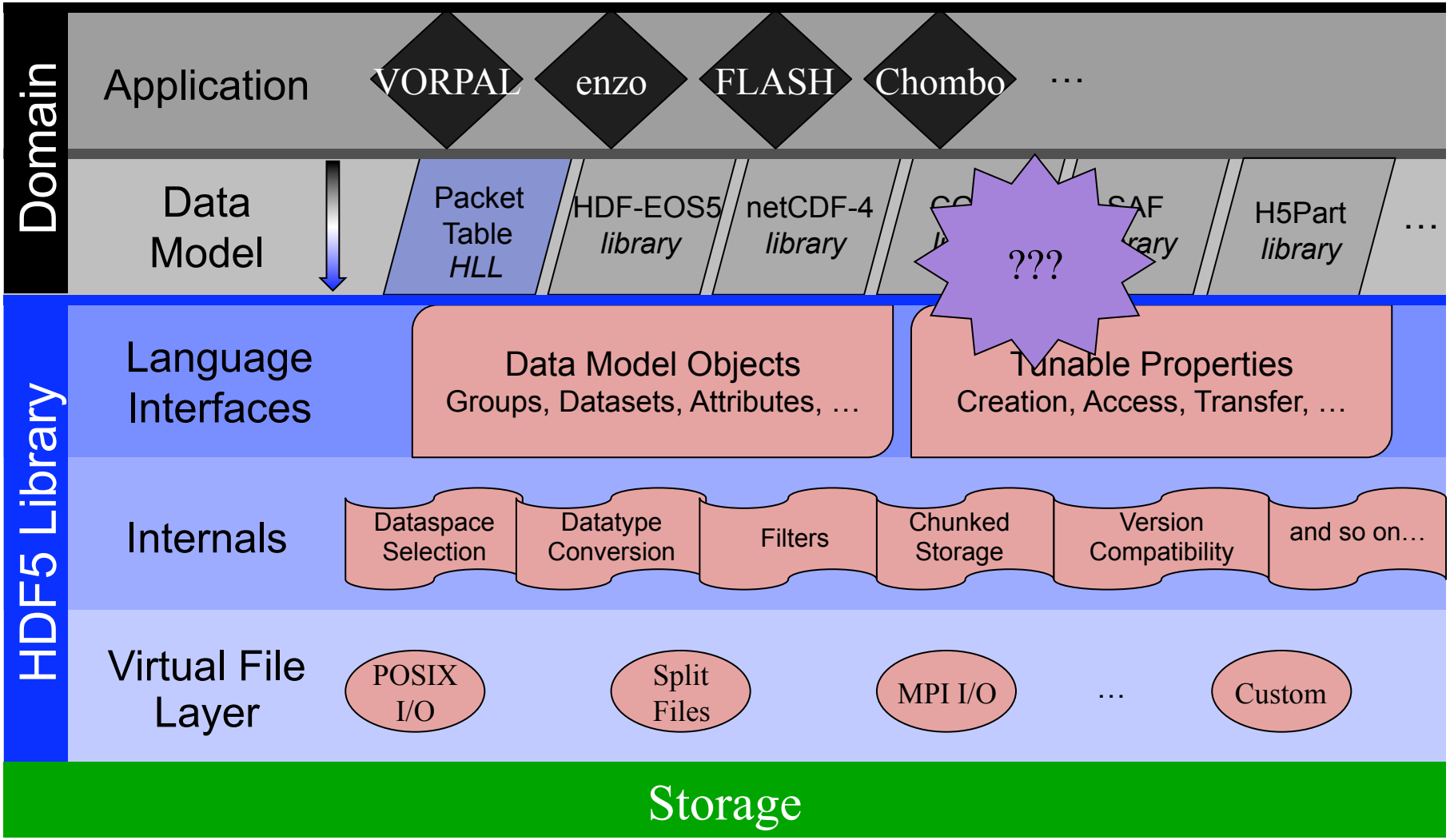
Data Model+HDF5 Layers



Application Design Decisions

- How does application instantiate the domain data model?
  - Domain-specific library (e.g., CNGS, HDF-EOS5) that calls HDF5 APIs *or* HDF5 APIs called directly?
  - Which of the HDF5 APIs will be used?
    - High-level library APIs
    - Java/Fortran/C++/C Language APIs
  - *Choices will affect performance*
    - *Layers of software between Application and Disk*
    - *Not all APIs expose the full set of tunable parameters*
    - *Inappropriate use of APIs/parameters can cause very bad performance*
    - *Appropriate parameters may vary from application to application and system to system*

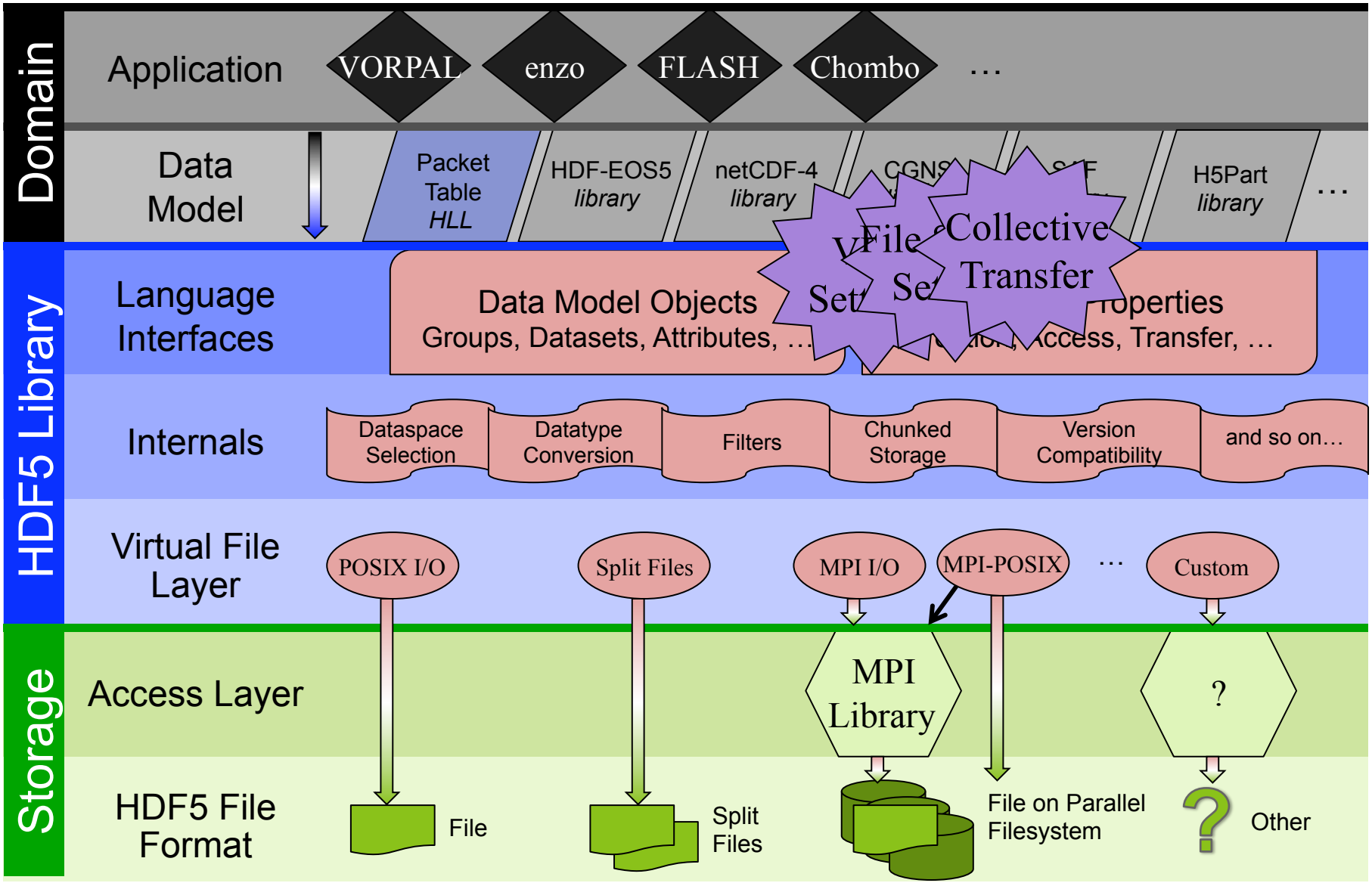
App+Data Model+HDF5 Layers



Storage Interface Decisions

- Which storage system is an application using?
  - Serial file system, parallel file system, network access, …
- How can the application take best advantage of it?
  - Serial applications vs. parallel applications
  - Parallel I/O Options
  - File System Options
  - *Again: Choices will affect performance!*
    - *More layers of software between Application and Disk*
    - *Not all APIs may expose the full set of tunable parameters*
    - *Inappropriate use of APIs/parameters can cause very bad performance*
    - *Appropriate parameters may vary from application to application and system to system*

App+Data Model+HDF5+Storage Layers



Tools For Identifying Bottlenecks

- Benchmarks:
  - Artificial: *h5perf,* IOR
  - More Realistic: I/O kernels (FLASH I/O, etc.)
- Performance Monitoring Tools:
  - Serial
    - Internal: *Log VFD in HDF5*
    - External: Traditional Profiling/Monitoring Tools, like gprof, Quantify, etc.
  - Parallel
    - Internal: MPI Profiling/Monitoring (MPE + Jumpshot)
    - External: TotalView (?)

Questions/Comments?